Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

Supporting Information for Paper

Non-Covalent Interactions Atlas Benchmark Data Sets 4: σ-Hole Interactions

Kristian Kříž and Jan Řezáč*

Institute of Organic Chemistry and Biochemistry, Czech Academy of Sciences, 166 10 Prague, Czech Republic

E-mail: rezac@uochb.cas.cz

May 31, 2022

Contents

1 Data Provided in Separate Files

The following ZIP archives are provided as a part of the Supporting Information:

- **ncia_sigmahole_SI_002.zip** contain geometries of all the systems as .xyz files, and tables of interaction energies and their components (tab-delimited text files). File names use the numbering defined in the paper. The headers of the .xyz files contain the definition of the two monomers in the complex, the benchmark interaction energy and selected other metadata.

- **ncia_sigmahole_SI_003.zip** contains structured, machine-readable but human friendly data file in YAML format (https://yaml.org/) defining and describing the data set. This file format is used by the Cuby framework (http://cuby4.molecular.cz), but it can be easily used on its own. The file contains all the metadata describing the systems, such as their assignment to groups. The last part of the file lists all the intermediate interaction energy components used to construct the benchmark, and results of some additional calculations discussed in the paper. (Please note that up to date version of the file will be bundled with the Cuby framework.)

Up to date version of all these files are also available at the GitHub repository `https://github.com/Honza-R/NCIAtlas`.

2 Lists of Monomers constituting the SH250 dataset

Table S1: Monomers used to construct the SH250 data set – compounds featuring a σ-hole.

Element	Molecule	Formula
Cl	Molecular Chlorine	Cl_2
	Chlorobenzene	C_6H_5Cl
	Chloromethane	CH_3Cl
	Trifluorochloromethane	CF_3Cl
Br	Molecular Bromine	Br_2
	Bromobenzene	C_6H_5Br
	Bromomethane	CH_3Br
	Trifluorobromomethane	CF_3Br
I	Molecular Iodine	I_2
	Iodobenzene	C_6H5I
	Iodomethane	CH_3I
	Trifluoroiodomethane	CF_3I
S	1,2-thiazole	C_3H_3NS
	1,2-thiazole-3-one	C_3H_3NSO
	Sulfurdifluoride	F_2S
	Carbonylsulfide	OCS
	Difluorodisulfane	F_2S_2
	Difluorothioformaldehyde	F_2CS
	Methylthiocyanate	CH_3SCN
	Thiophene-3(2H)-one	C_4H_4OS
Se	1,2-selenazole	C_3H_3NSe
	1,2-selenazole-3-one	C_3H_3NSeO
	Seleniumdifluoride	F_2Se
	Carbonylselenide	OCSe
	Difluorodiselane	F_2Se_2
	Difluoroselenoformaldehyde	F_2CSe
	Methylselenocyanate	CH_3SeCN
	Selenophene-3(2H)-one	C_4H_4OSe
	Tetrafluoroselenourea	$F4N_2CSe$
P	Bis(difluorophosphanyl)thioether	F_4P_2S
	Fluorobromocyanophosphine	NCFBrP
	Fluorophosphine	H_2FP
	Perfluorodimethyl cyanophosphine	F_6C_3NP
	Tricyanophosphine	N_3C_3P
As	Bis(difluoroarsanyl)thioether	F_4As_2S
	Fluorobromocyanoarsine	NCFBrAs
	Fluoroarsine	H_2FAs
	Perfluorodimethyl cyanoarsine	F_6C_3NAs
	Tricyanoarsine	N_3C_3As

Table S2: Monomers used to construct the SH250 data set – electron donors

Element	Molecule	Formula
		Electron Donors
O	Acetone	C_3H_6O
	Furan	H_4C_4O
	Nitromethane	CH_3NO_2
	Pyridine-N-oxide	C_5H_5NO
N	Pyrazine	$C_4H_4N_2$
	Trimethylamine	C_3H_9N
	Acetonitrile	H_3C_2N
S	Dimethylthioether	C_2H_6S
	Thioacetone	C_3H_6S
F	Molecular Fluorine	F_2
Cl	Molecular Chlorine	Cl_2
Br	Molecular Bromine	Br_2
I	Molecular Iodine	I_2
C	Benzene	C_6H_6

Table S3: The number of complexes representing each element pair. For each element pair, there are 10 separation points in the SH250x10 dataset.

σ-hole	Electron donors							
element	O	N	C	S	F	Cl	Br	I
Cl	10	8	2	5	1	1	1	1
Br	14	7	3	5	1	2	2	2
I	15	8	3	6	1	3	3	3
S	15	16	0	0	0	0	0	0
Se	22	22	0	0	0	0	0	0
P	19	14	0	0	0	0	0	0
As	20	15	0	0	0	0	0	0

3 Predefined Subsets of the SH250 dataset

Table S4: Assignment of the systems to the smaller subsets obtained by clustering analysis (20, 50, 100 and 200 systems). (Note that this information is provided also in the data files.)

Subset size	Systems
20	1.1.01, 1.2.02, 1.2.08, 2.2.02, 2.3.01, 2.3.10, 2.7.01, 3.1.02, 3.3.10, 3.3.15, 3.4.03, 4.1.11, 4.2.05, 5.1.05, 5.1.11, 6.1.11, 6.2.04, 6.2.16, 7.2.01, 7.2.03
50	1.2.07, 1.2.08, 1.3.01, 1.3.02, 1.3.03, 1.3.07, 1.4.03, 1.5.01, 1.7.01, 2.1.01, 2.2.01, 2.2.06, 2.3.04, 2.3.05, 2.3.08, 2.3.10, 2.5.01, 3.1.02, 3.2.01, 3.2.06, 3.3.04, 3.3.10, 3.3.11, 3.4.01, 3.4.03, 3.6.01, 3.7.01, 3.7.02, 3.8.02, 4.1.09, 4.1.15, 4.2.04, 4.2.07, 4.2.09, 4.2.10, 4.2.15, 5.1.05, 5.1.09, 5.2.01, 6.1.03, 6.1.09, 6.1.11, 6.2.03, 6.2.14, 6.2.16, 6.2.18, 7.1.03, 7.2.01, 7.2.03, 7.2.12
100	1.1.01, 1.1.02, 1.2.03, 1.2.08, 1.3.01, 1.3.02, 1.3.03, 1.3.07, 1.3.09, 1.4.01, 1.4.03, 1.5.01, 1.6.01, 1.8.01, 2.1.01, 2.2.01, 2.2.03, 2.2.06, 2.3.01, 2.3.02, 2.3.03, 2.3.05, 2.3.06, 2.3.09, 2.3.10, 2.3.11, 2.3.13, 2.3.14, 2.4.01, 2.4.04, 2.5.01, 2.7.01, 2.7.02, 3.1.02, 3.2.01, 3.2.02, 3.2.03, 3.2.04, 3.2.05, 3.3.02, 3.3.03, 3.3.04, 3.3.06, 3.3.07, 3.3.14, 3.4.01, 3.4.02, 3.4.03, 3.5.01, 3.6.01, 3.6.03, 3.7.01, 3.7.02, 3.7.03, 3.8.02, 4.1.01, 4.1.03, 4.1.05, 4.1.09, 4.1.15, 4.2.01, 4.2.02, 4.2.04, 4.2.07, 4.2.08, 4.2.11, 4.2.13, 4.2.14, 5.1.01, 5.1.05, 5.1.07, 5.1.15, 5.1.16, 5.1.20, 5.2.09, 5.2.10, 5.2.15, 6.1.03, 6.1.05, 6.1.06, 6.1.07, 6.1.09, 6.1.13, 6.2.03, 6.2.04, 6.2.06, 6.2.09, 6.2.12, 6.2.13, 6.2.16, 6.2.17, 6.2.18, 7.1.03, 7.2.01, 7.2.07, 7.2.11, 7.2.13, 7.2.17, 7.2.18, 7.2.19
200	1.1.01, 1.1.02, 1.2.01, 1.2.02, 1.2.03, 1.2.06, 1.2.08, 1.3.01, 1.3.02, 1.3.03, 1.3.04, 1.3.05, 1.3.06, 1.3.07, 1.3.08, 1.3.09, 1.4.01, 1.4.02, 1.4.03, 1.5.01, 1.6.01, 1.7.01, 1.8.01, 2.1.01, 2.1.02, 2.1.03, 2.2.01, 2.2.02, 2.2.03, 2.2.05, 2.2.06, 2.3.01, 2.3.02, 2.3.03, 2.3.05, 2.3.06, 2.3.07, 2.3.09, 2.3.10, 2.3.11, 2.3.12, 2.3.13, 2.3.14, 2.4.01, 2.4.02, 2.4.05, 2.5.01, 2.6.01, 2.6.02, 2.7.01, 2.7.02, 2.8.01, 2.8.02, 3.1.01, 3.1.02, 3.1.03, 3.2.01, 3.2.02, 3.2.03, 3.2.04, 3.2.05, 3.2.07, 3.2.08, 3.3.01, 3.3.02, 3.3.04, 3.3.05, 3.3.06, 3.3.08, 3.3.09, 3.3.10, 3.3.12, 3.3.13, 3.3.14, 3.3.15, 3.4.01, 3.4.02, 3.4.03, 3.4.04, 3.4.05, 3.5.01, 3.6.01, 3.6.02, 3.6.03, 3.7.01, 3.7.03, 3.8.01, 3.8.02, 3.8.03, 4.1.01, 4.1.02, 4.1.03, 4.1.04, 4.1.05, 4.1.06, 4.1.07, 4.1.08, 4.1.09, 4.1.10, 4.1.11, 4.1.12, 4.1.13, 4.1.14, 4.1.15, 4.2.01, 4.2.02, 4.2.03, 4.2.04, 4.2.05, 4.2.06, 4.2.08, 4.2.09, 4.2.10, 4.2.11, 4.2.12, 4.2.13, 4.2.14, 4.2.15, 5.1.01, 5.1.02, 5.1.03, 5.1.05, 5.1.06, 5.1.07, 5.1.08, 5.1.13, 5.1.14, 5.1.15, 5.1.17, 5.1.21, 5.2.01, 5.2.02, 5.2.03, 5.2.04, 5.2.05, 5.2.06, 5.2.07, 5.2.08, 5.2.10, 5.2.11, 5.2.12, 5.2.15, 5.2.17, 5.2.18, 5.2.19, 5.2.21, 5.2.22, 6.1.01, 6.1.02, 6.1.03, 6.1.05, 6.1.06, 6.1.07, 6.1.08, 6.1.09, 6.1.10, 6.1.11, 6.1.12, 6.1.13, 6.1.14, 6.2.01, 6.2.02, 6.2.03, 6.2.04, 6.2.05, 6.2.06, 6.2.07, 6.2.08, 6.2.09, 6.2.11, 6.2.12, 6.2.13, 6.2.14, 6.2.16, 6.2.17, 6.2.18, 6.2.19, 7.1.01, 7.1.03, 7.1.04, 7.1.07, 7.1.09, 7.1.12, 7.1.13, 7.1.14, 7.2.01, 7.2.02, 7.2.04, 7.2.05, 7.2.06, 7.2.07, 7.2.08, 7.2.11, 7.2.14, 7.2.15, 7.2.16, 7.2.17, 7.2.18, 7.2.19, 7.2.20

4 DFT Results

Figure S1: Systematic error (MSE) of DFT methods in the SH250 data set (equilibrium geometries only) and its groups by the interacting element. The functionals are ordered by increasing RMSE in the whole data set. The source data are listed in Table S7.

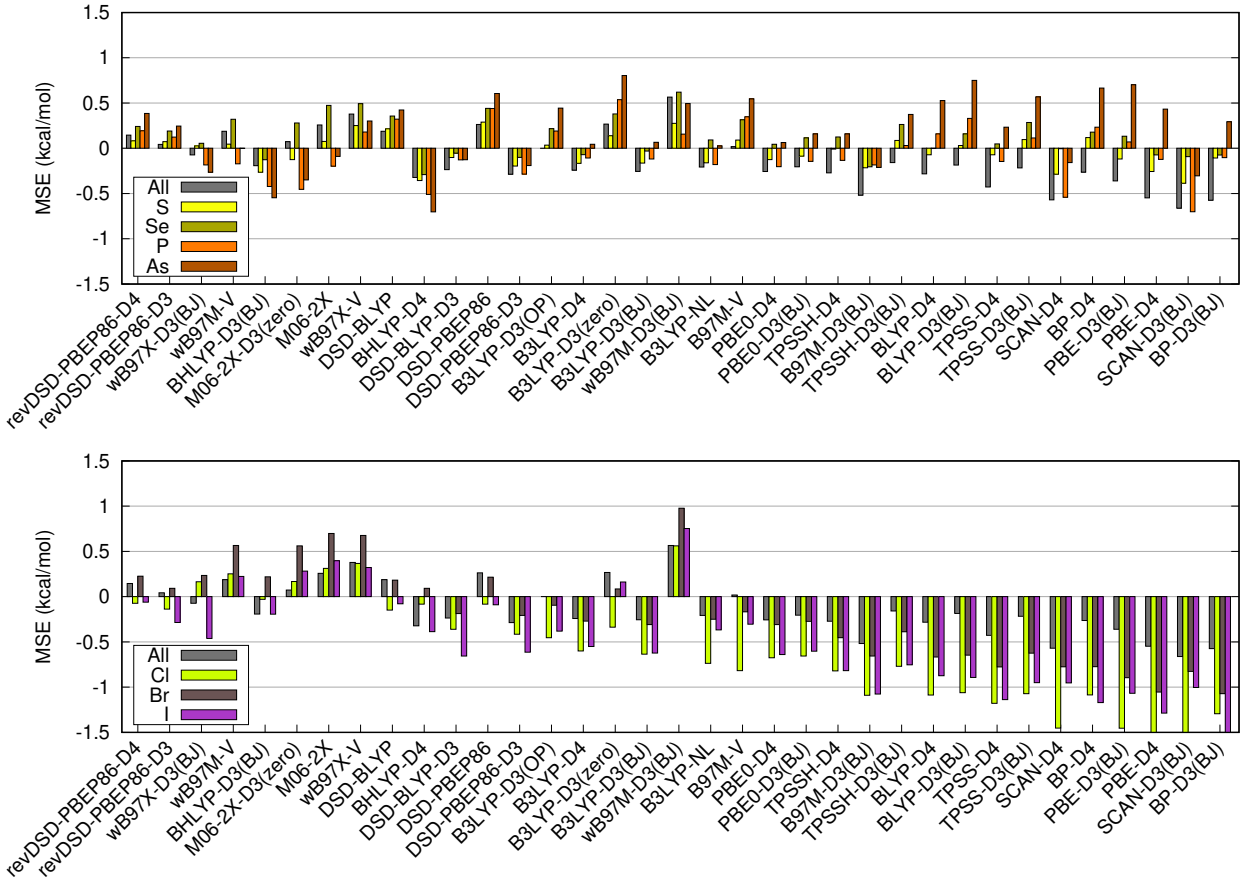

Figure S2: RMSE of DFT methods in the SH250×10 data set and its groups by the interacting element. The DFT functional labels are coloured by the rung of the functional: double-hybrids(red), range-separated hybrids (blue), hybrids (black) and GGA and meta-GGA (green). The functionals are ordered by increasing RMSE in the whole data set. The source data are listed in Table S6.

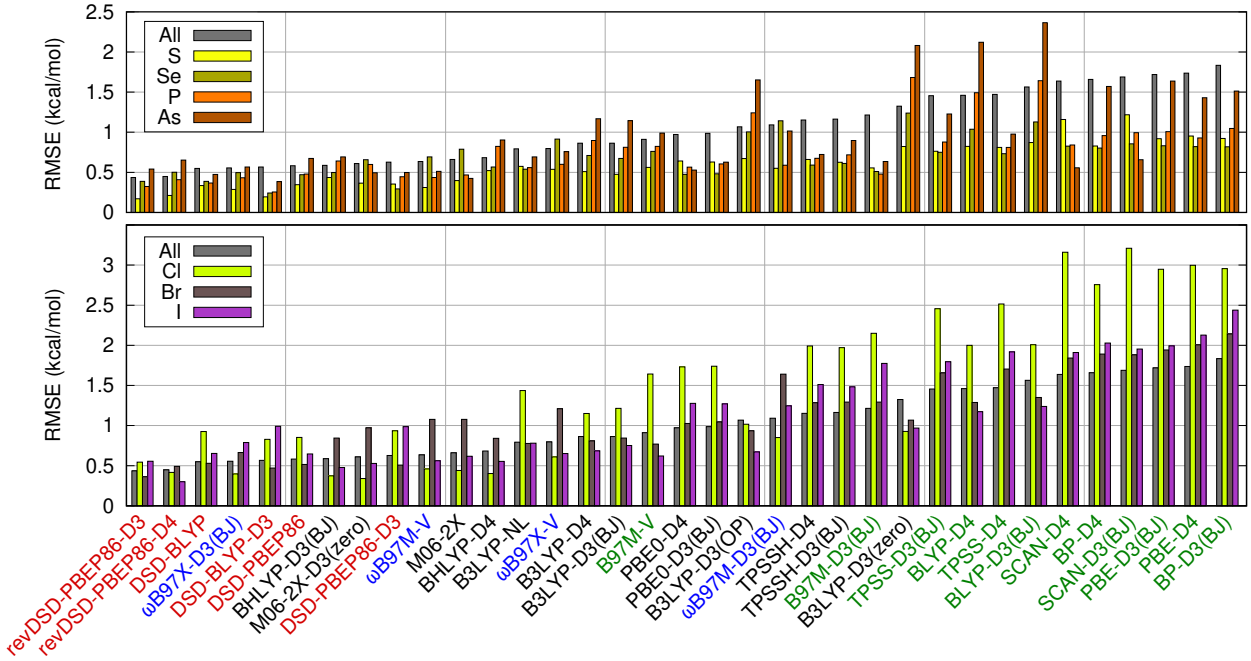

Figure S3: Systematic error (MSE) of DFT methods in the SH250×10 data set and its groups by the interacting element. The functionals are ordered by increasing RMSE in the whole data set. The source data are listed in Table S8.

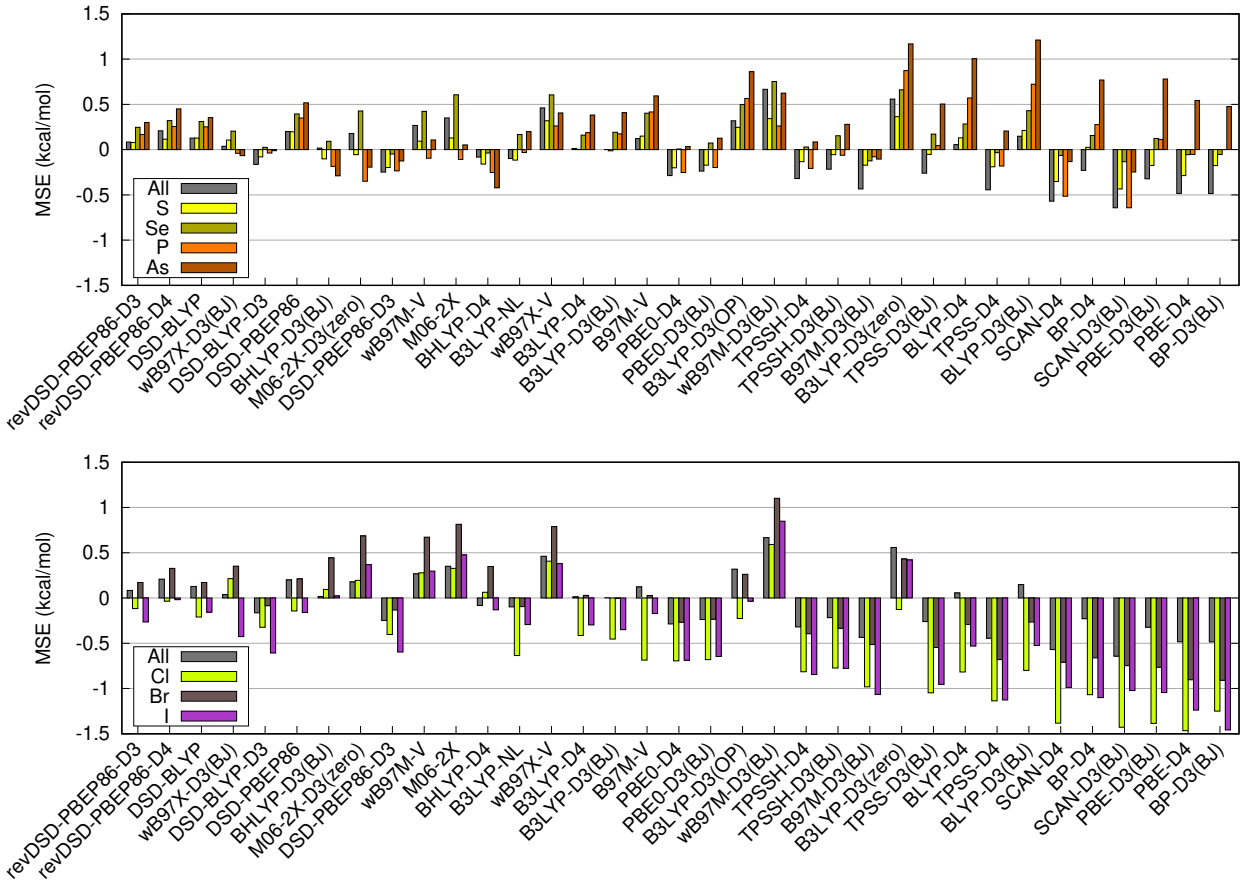

Table S5: RMSE (in kcal/mol) of DFT calculations in the SH250 data set (250 equilibrium geometries) and its groups by the electron acceptor element. The functionals are ordered by increasing RMSE in the whole data set.

Method	All	Cl	Br	I	S	Se	P	As
revDSD-PBEP86-D4	0.338	0.458	0.303	0.294	0.152	0.306	0.274	0.484
revDSD-PBEP86-D3	0.397	0.582	0.349	0.531	0.165	0.263	0.264	0.454
wB97X-D3(BJ)	0.423	0.300	0.417	0.628	0.168	0.262	0.401	0.536
wB97M-V	0.431	0.347	0.744	0.337	0.213	0.418	0.374	0.359
BHLYP-D3(BJ)	0.448	0.291	0.380	0.335	0.397	0.263	0.602	0.715
M06-2X-D3(zero)	0.467	0.299	0.635	0.356	0.379	0.408	0.586	0.505
M06-2X	0.495	0.408	0.764	0.453	0.361	0.542	0.406	0.354
wB97X-V	0.522	0.457	0.860	0.424	0.308	0.566	0.362	0.435
DSD-BLYP	0.525	0.837	0.570	0.536	0.353	0.401	0.392	0.504
BHLYP-D4	0.568	0.299	0.382	0.480	0.491	0.415	0.754	0.916
DSD-BLYP-D3	0.569	0.859	0.579	0.922	0.216	0.224	0.275	0.399
DSD-PBEP86	0.570	0.778	0.546	0.539	0.383	0.476	0.514	0.707
DSD-PBEP86-D3	0.607	0.906	0.593	0.880	0.326	0.293	0.458	0.512
B3LYP-D3(OP)	0.657	1.224	0.626	0.682	0.270	0.352	0.411	0.708
B3LYP-D4	0.677	1.361	0.721	0.782	0.334	0.282	0.327	0.414
B3LYP-D3(zero)	0.679	1.085	0.503	0.424	0.297	0.492	0.716	0.989
B3LYP-D3(BJ)	0.751	1.446	0.828	0.922	0.354	0.269	0.362	0.496
wB97M-D3(BJ)	0.751	0.658	1.181	0.865	0.346	0.740	0.369	0.642
B3LYP-NL	0.792	1.617	0.934	0.844	0.531	0.339	0.302	0.285
B97M-V	0.821	1.769	0.732	0.586	0.348	0.457	0.583	0.754
PBE0-D4	0.897	1.672	1.076	1.134	0.504	0.355	0.411	0.396
PBE0-D3(BJ)	0.929	1.690	1.102	1.151	0.517	0.391	0.501	0.545
TPSSH-D4	1.158	2.107	1.449	1.459	0.564	0.474	0.527	0.603
B97M-D3(BJ)	1.168	2.217	1.357	1.482	0.497	0.439	0.500	0.669
TPSSH-D3(BJ)	1.191	2.100	1.462	1.468	0.570	0.535	0.635	0.819
BLYP-D4	1.206	2.315	1.430	1.318	0.472	0.444	0.584	0.996
BLYP-D3(BJ)	1.283	2.334	1.489	1.420	0.474	0.463	0.744	1.248
TPSS-D4	1.485	2.678	1.902	1.874	0.683	0.571	0.659	0.797
TPSS-D3(BJ)	1.487	2.632	1.862	1.791	0.683	0.634	0.774	1.075
SCAN-D4	1.614	3.210	1.961	1.767	1.019	0.709	0.765	0.522
BP-D4	1.617	2.869	2.065	1.991	0.708	0.624	0.752	1.118
PBE-D3(BJ)	1.646	3.017	2.074	1.899	0.762	0.600	0.787	1.226
PBE-D4	1.661	3.058	2.135	2.018	0.777	0.581	0.685	0.987
SCAN-D3(BJ)	1.669	3.263	2.005	1.812	1.081	0.745	0.947	0.654
BP-D3(BJ)	1.849	3.077	2.343	2.438	0.833	0.740	1.029	1.224

Table S6: RMSE (in kcal/mol) of DFT calculations in the SH250×10 data set and its groups by the electron acceptor element. The functionals are ordered by increasing RMSE in the whole data set.

Method	All	Cl	Br	I	S	Se	P	As
revDSD-PBEP86-D3	0.436	0.545	0.363	0.555	0.171	0.388	0.324	0.541
revDSD-PBEP86-D4	0.449	0.417	0.492	0.300	0.212	0.501	0.408	0.652
DSD-BLYP	0.549	0.926	0.530	0.653	0.335	0.388	0.367	0.474
wB97X-D3(BJ)	0.556	0.397	0.664	0.788	0.287	0.496	0.432	0.566
DSD-BLYP-D3	0.567	0.829	0.471	0.991	0.195	0.242	0.256	0.385
DSD-PBEP86	0.583	0.852	0.515	0.647	0.347	0.471	0.478	0.673
BHLYP-D3(BJ)	0.588	0.374	0.844	0.477	0.435	0.498	0.641	0.693
M06-2X-D3(zero)	0.609	0.340	0.972	0.529	0.365	0.658	0.598	0.495
DSD-PBEP86-D3	0.628	0.936	0.507	0.986	0.354	0.294	0.446	0.498
wB97M-V	0.635	0.460	1.078	0.561	0.311	0.692	0.436	0.512
M06-2X	0.661	0.440	1.077	0.618	0.397	0.788	0.467	0.425
BHLYP-D4	0.683	0.402	0.841	0.554	0.522	0.566	0.823	0.904
B3LYP-NL	0.793	1.437	0.778	0.780	0.574	0.539	0.558	0.693
wB97X-V	0.797	0.609	1.209	0.651	0.538	0.915	0.600	0.758
B3LYP-D4	0.863	1.150	0.809	0.684	0.509	0.710	0.897	1.168
B3LYP-D3(BJ)	0.864	1.215	0.844	0.752	0.475	0.674	0.813	1.144
B97M-V	0.911	1.642	0.769	0.621	0.561	0.761	0.823	0.990
PBE0-D4	0.972	1.731	1.027	1.276	0.641	0.476	0.565	0.528
PBE0-D3(BJ)	0.987	1.739	1.046	1.271	0.629	0.482	0.605	0.628
B3LYP-D3(OP)	1.068	1.017	0.936	0.672	0.672	1.004	1.241	1.652
wB97M-D3(BJ)	1.091	0.850	1.641	1.247	0.551	1.143	0.589	1.015
TPSSH-D4	1.153	1.991	1.285	1.512	0.661	0.591	0.675	0.723
TPSSH-D3(BJ)	1.164	1.971	1.292	1.482	0.625	0.612	0.718	0.897
B97M-D3(BJ)	1.215	2.149	1.292	1.775	0.556	0.510	0.478	0.635
B3LYP-D3(zero)	1.325	0.927	1.068	0.969	0.822	1.238	1.682	2.081
TPSS-D3(BJ)	1.456	2.456	1.656	1.796	0.763	0.750	0.878	1.227
BLYP-D4	1.460	1.999	1.288	1.173	0.823	1.038	1.493	2.121
TPSS-D4	1.472	2.516	1.704	1.919	0.811	0.732	0.811	0.977
BLYP-D3(BJ)	1.564	2.008	1.351	1.239	0.872	1.129	1.643	2.365
SCAN-D4	1.638	3.160	1.841	1.911	1.159	0.827	0.841	0.557
BP-D4	1.647	2.756	1.891	2.028	0.828	0.802	0.957	1.481
SCAN-D3(BJ)	1.688	3.210	1.881	1.953	1.217	0.855	0.995	0.657
PBE-D3(BJ)	1.719	2.948	1.942	1.993	0.919	0.832	1.007	1.638
PBE-D4	1.736	2.998	2.006	2.128	0.953	0.820	0.929	1.430
BP-D3(BJ)	1.822	2.956	2.142	2.438	0.923	0.819	1.047	1.414

Table S7: Systematic error (MSE, in kcal/mol) of DFT calculations in the SH250 data set (250 equilibrium geometries) and its groups by the electron acceptor element. The functionals are ordered by increasing RMSE in the whole data set.

Method	All	Cl	Br	I	S	Se	P	As
revDSD-PBEP86-D4	0.146	-0.073	0.227	-0.061	0.083	0.241	0.194	0.385
revDSD-PBEP86-D3	0.043	-0.137	0.092	-0.285	0.073	0.191	0.124	0.245
wB97X-D3(BJ)	-0.073	0.165	0.235	-0.462	0.026	0.056	-0.184	-0.265
wB97M-V	0.188	0.253	0.565	0.223	0.046	0.321	-0.171	-0.001
BHLYP-D3(BJ)	-0.192	-0.030	0.220	-0.193	-0.265	-0.126	-0.422	-0.547
M06-2X-D3(zero)	0.073	0.168	0.561	0.283	-0.125	0.280	-0.453	-0.349
M06-2X	0.258	0.313	0.699	0.398	0.076	0.474	-0.198	-0.089
wB97X-V	0.378	0.367	0.677	0.322	0.251	0.493	0.180	0.301
DSD-BLYP	0.188	-0.149	0.183	-0.078	0.216	0.357	0.322	0.423
BHLYP-D4	-0.322	-0.083	0.092	-0.386	-0.357	-0.291	-0.509	-0.702
DSD-BLYP-D3	-0.236	-0.360	-0.187	-0.656	-0.102	-0.054	-0.128	-0.127
DSD-PBEP86	0.263	-0.083	0.216	-0.090	0.289	0.441	0.440	0.605
DSD-PBEP86-D3	-0.287	-0.415	-0.207	-0.613	-0.196	-0.100	-0.285	-0.191
B3LYP-D3(OP)	-0.000	-0.454	-0.095	-0.381	0.035	0.217	0.191	0.444
B3LYP-D4	-0.242	-0.600	-0.270	-0.551	-0.166	-0.070	-0.107	0.044
B3LYP-D3(zero)	0.268	-0.337	0.086	0.162	0.139	0.380	0.537	0.804
B3LYP-D3(BJ)	-0.255	-0.635	-0.308	-0.623	-0.164	-0.031	-0.119	0.066
wB97M-D3(BJ)	0.566	0.561	0.978	0.753	0.276	0.621	0.156	0.494
B3LYP-NL	-0.207	-0.736	-0.249	-0.367	-0.161	0.092	-0.180	0.026
B97M-V	0.019	-0.817	-0.167	-0.304	0.091	0.316	0.348	0.547
PBE0-D4	-0.256	-0.675	-0.308	-0.640	-0.125	0.045	-0.203	0.063
PBE0-D3(BJ)	-0.204	-0.656	-0.273	-0.602	-0.087	0.116	-0.145	0.160
TPSSH-D4	-0.272	-0.820	-0.453	-0.817	-0.009	0.125	-0.134	0.159
B97M-D3(BJ)	-0.518	-1.090	-0.656	-1.076	-0.216	-0.204	-0.184	-0.211
TPSSH-D3(BJ)	-0.158	-0.770	-0.388	-0.753	0.087	0.264	0.031	0.374
BLYP-D4	-0.283	-1.087	-0.667	-0.874	-0.072	-0.001	0.160	0.528
BLYP-D3(BJ)	-0.185	-1.061	-0.646	-0.893	0.031	0.160	0.330	0.751
TPSS-D4	-0.427	-1.179	-0.777	-1.138	-0.072	0.049	-0.146	0.233
TPSS-D3(BJ)	-0.217	-1.072	-0.624	-0.950	0.096	0.285	0.114	0.569
SCAN-D4	-0.570	-1.452	-0.775	-0.953	-0.286	-0.005	-0.541	-0.158
BP-D4	-0.264	-1.085	-0.774	-1.170	0.118	0.179	0.233	0.665
PBE-D3(BJ)	-0.360	-1.453	-0.895	-1.068	-0.118	0.134	0.071	0.702
PBE-D4	-0.548	-1.552	-1.054	-1.286	-0.257	-0.073	-0.123	0.433
SCAN-D3(BJ)	-0.662	-1.508	-0.826	-1.004	-0.387	-0.092	-0.701	-0.305
BP-D3(BJ)	-0.574	-1.293	-1.072	-1.615	-0.107	-0.074	-0.103	0.294

Table S8: Systematic error (MSE, in kcal/mol) of DFT calculations in the SH250×10 data set and its groups by the electron acceptor element. The functionals are ordered by increasing RMSE in the whole data set.

Method	All	Cl	Br	I	S	Se	P	As
revDSD-PBEP86-D3	0.084	-0.117	0.172	-0.265	0.078	0.247	0.167	0.300
revDSD-PBEP86-D4	0.208	-0.035	0.327	-0.018	0.115	0.321	0.256	0.451
DSD-BLYP	0.128	-0.210	0.172	-0.156	0.128	0.311	0.252	0.355
wB97X-D3(BJ)	0.038	0.214	0.352	-0.426	0.106	0.204	-0.040	-0.067
DSD-BLYP-D3	-0.164	-0.324	-0.086	-0.608	-0.080	0.025	-0.038	-0.011
DSD-PBEP86	0.200	-0.143	0.211	-0.160	0.198	0.394	0.350	0.518
BHLYP-D3(BJ)	0.017	0.095	0.444	0.023	-0.102	0.092	-0.185	-0.289
M06-2X-D3(zero)	0.178	0.195	0.686	0.368	-0.056	0.427	-0.350	-0.194
DSD-PBEP86-D3	-0.248	-0.403	-0.133	-0.597	-0.197	-0.049	-0.234	-0.125
wB97M-V	0.268	0.278	0.672	0.297	0.094	0.424	-0.095	0.108
M06-2X	0.351	0.327	0.814	0.477	0.129	0.607	-0.109	0.052
BHLYP-D4	-0.083	0.063	0.348	-0.131	-0.160	-0.038	-0.252	-0.422
B3LYP-NL	-0.098	-0.635	-0.095	-0.294	-0.115	0.168	-0.032	0.199
wB97X-V	0.462	0.407	0.790	0.381	0.319	0.605	0.262	0.405
B3LYP-D4	0.013	-0.414	0.029	-0.297	0.003	0.161	0.188	0.383
B3LYP-D3(BJ)	0.001	-0.454	-0.004	-0.349	-0.013	0.192	0.175	0.409
B97M-V	0.123	-0.686	0.028	-0.171	0.150	0.403	0.416	0.594
PBE0-D4	-0.286	-0.694	-0.268	-0.688	-0.200	0.008	-0.252	0.035
PBE0-D3(BJ)	-0.238	-0.680	-0.235	-0.646	-0.171	0.073	-0.198	0.127
B3LYP-D3(OP)	0.319	-0.227	0.262	-0.035	0.247	0.499	0.565	0.861
wB97M-D3(BJ)	0.667	0.592	1.102	0.849	0.342	0.753	0.262	0.625
TPSSH-D4	-0.320	-0.814	-0.395	-0.845	-0.133	0.029	-0.207	0.086
TPSSH-D3(BJ)	-0.217	-0.773	-0.334	-0.778	-0.055	0.154	-0.062	0.280
B97M-D3(BJ)	-0.434	-0.981	-0.514	-1.065	-0.171	-0.124	-0.078	-0.104
B3LYP-D3(zero)	0.559	-0.126	0.432	0.420	0.365	0.661	0.873	1.169
TPSS-D3(BJ)	-0.260	-1.047	-0.546	-0.954	-0.052	0.172	0.045	0.504
BLYP-D4	0.056	-0.817	-0.292	-0.531	0.131	0.284	0.571	1.005
TPSS-D4	-0.444	-1.136	-0.681	-1.127	-0.190	-0.034	-0.181	0.206
BLYP-D3(BJ)	0.148	-0.799	-0.266	-0.525	0.212	0.430	0.723	1.212
SCAN-D4	-0.570	-1.382	-0.709	-0.988	-0.352	-0.065	-0.515	-0.132
BP-D4	-0.231	-1.068	-0.661	-1.100	0.025	0.157	0.277	0.751
SCAN-D3(BJ)	-0.642	-1.428	-0.748	-1.023	-0.434	-0.133	-0.642	-0.247
PBE-D3(BJ)	-0.322	-1.385	-0.765	-1.045	-0.175	0.124	0.113	0.780
PBE-D4	-0.484	-1.466	-0.903	-1.239	-0.287	-0.055	-0.052	0.544
BP-D3(BJ)	-0.488	-1.249	-0.909	-1.458	-0.176	-0.054	0.001	0.457

5 SQM Results

Table S9: RMSE of SQM methods tested in the SH250 data set (equilibrium geometries only). All values in kcal/mol.

Method	All	Cl	Br	I	S	Se	P	As
PM6	5.873	1.532	5.631	3.751	2.203	7.409	3.684	10.433
PM6-D3	5.367	1.221	5.933	4.029	1.324	6.686	2.556	9.216
PM6-D3H4	5.367	1.221	5.933	4.029	1.324	6.686	2.556	9.216
PM6-D3H4X	10.279	26.312	4.129	3.544	1.318	6.680	2.549	9.206
PM6-D3H4X2	4.705	1.256	2.138	1.724	1.318	6.680	2.549	9.207
PM7	4.273	1.004	1.655	6.398	2.856	4.958	4.452	4.566
GFN2-XTB	3.014	0.905	1.287	1.538	1.518	2.403	3.825	6.046
GFN-XTB	3.599	1.720	2.217	1.722	2.434	2.349	4.919	6.816

Table S10: Systematic error (MSE) of SQM methods tested in the SH250 data set (equilibrium geometries only). All values in kcal/mol.

Method	All	Cl	Br	I	S	Se	P	As
PM6	2.606	0.986	-2.089	-1.844	1.972	6.253	3.201	9.531
PM6-D3	1.780	0.423	-2.655	-2.332	1.116	5.432	1.951	8.236
PM6-D3H4	1.780	0.423	-2.655	-2.332	1.116	5.432	1.951	8.236
PM6-D3H4X	3.825	8.891	1.196	0.720	1.109	5.425	1.942	8.225
PM6-D3H4X2	2.724	0.855	0.783	0.069	1.109	5.425	1.943	8.225
PM7	2.702	-0.293	0.976	4.734	2.388	3.978	2.348	3.527
GFN2-XTB	-1.267	0.548	0.754	0.302	0.386	-2.066	-3.249	-5.321
GFN-XTB	-0.865	0.861	0.690	-0.011	0.846	1.042	-4.001	-5.875

Table S11: RMSE of SQM methods tested in the closest points of the SH250×10 dissociation curves. One system (2.7.01, bromine dimer) had to be excluded because of convergence problems. All values in kcal/mol.

Method	All	Cl	Br	I	S	Se	P	As
PM6	11.716	6.600	17.184	17.809	5.901	6.776	5.898	11.778
PM6-D3	11.830	6.995	17.638	18.215	6.459	6.423	5.977	10.889
PM6-D3H4	11.830	6.995	17.638	18.215	6.459	6.423	5.977	10.889
PM6-D3H4X	19.677	49.583	13.069	16.124	6.471	6.418	5.989	10.879
PM6-D3H4X2	9.874	6.387	12.305	14.882	6.471	6.418	5.989	10.880
PM7	11.211	4.857	4.567	14.193	2.455	5.725	23.180	7.864
GFN2-XTB	5.918	2.719	2.786	3.146	5.276	6.323	9.812	7.598
GFN-XTB	7.606	4.264	3.480	5.905	6.709	7.184	12.424	9.674

Table S12: MSE of SQM methods tested in the closest points of the SH250×10 dissociation curves. One system (2.7.01, bromine dimer) had to be excluded because of convergence problems. All values in kcal/mol.

Method	All	Cl	Br	I	S	Se	P	As
PM6	-3.530	-4.220	-12.432	-15.043	-4.799	3.596	-0.031	8.878
PM6-D3	-4.271	-4.794	-12.965	-15.474	-5.528	2.913	-1.169	7.702
PM6-D3H4	-4.271	-4.794	-12.965	-15.474	-5.528	2.913	-1.169	7.702
PM6-D3H4X	2.816	27.645	-4.282	-3.061	-5.541	2.896	-1.190	7.678
PM6-D3H4X2	-1.634	-4.186	-5.270	-6.728	-5.541	2.896	-1.189	7.679
PM7	1.253	-3.584	-0.218	11.193	-0.179	-0.962	1.025	-0.889
GFN2-XTB	-3.428	-1.663	0.553	1.286	-3.995	-5.759	-9.143	-5.708
GFN-XTB	-5.256	-2.868	-1.618	-3.874	-4.223	-4.953	-11.506	-8.035

6 SQM Results - dissociation curve plots

Figure S4: SH250×10 dissociation curves computed with semiempirical QM methods - chlorine

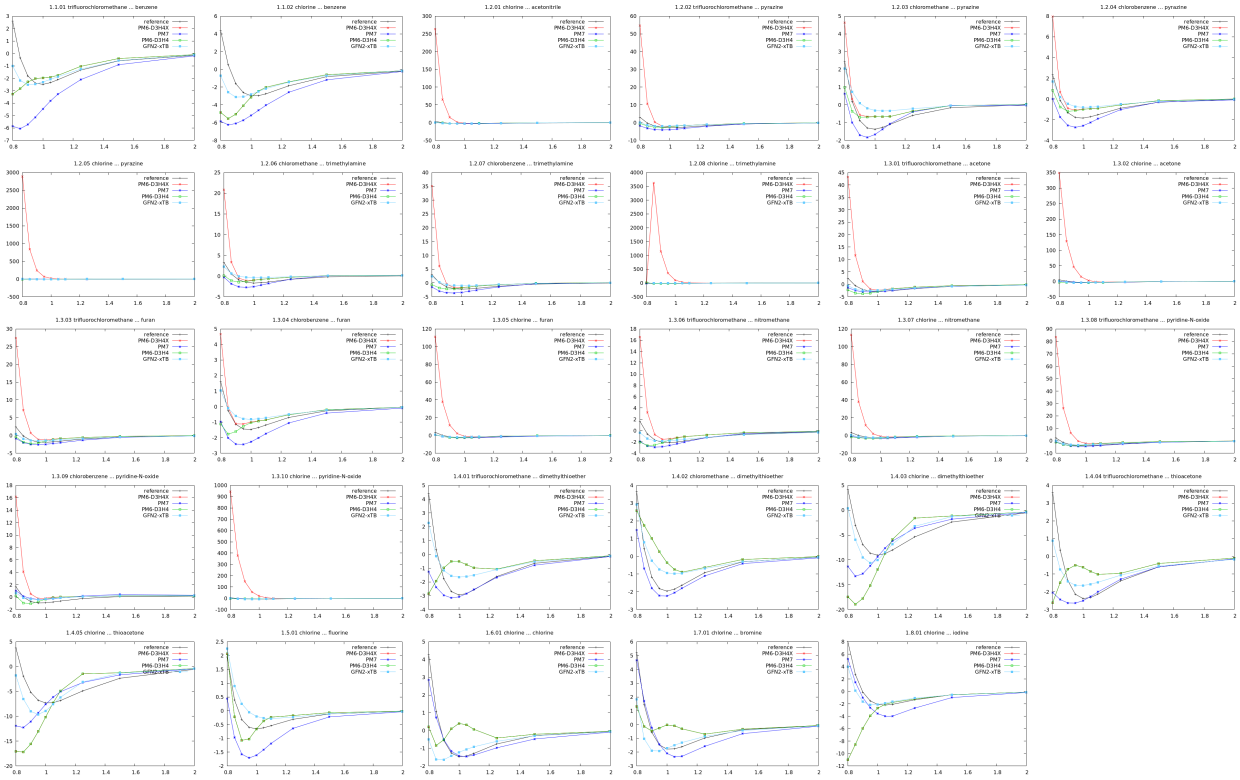

Figure S5: SH250×10 dissociation curves computed with semiempirical QM methods - bromine

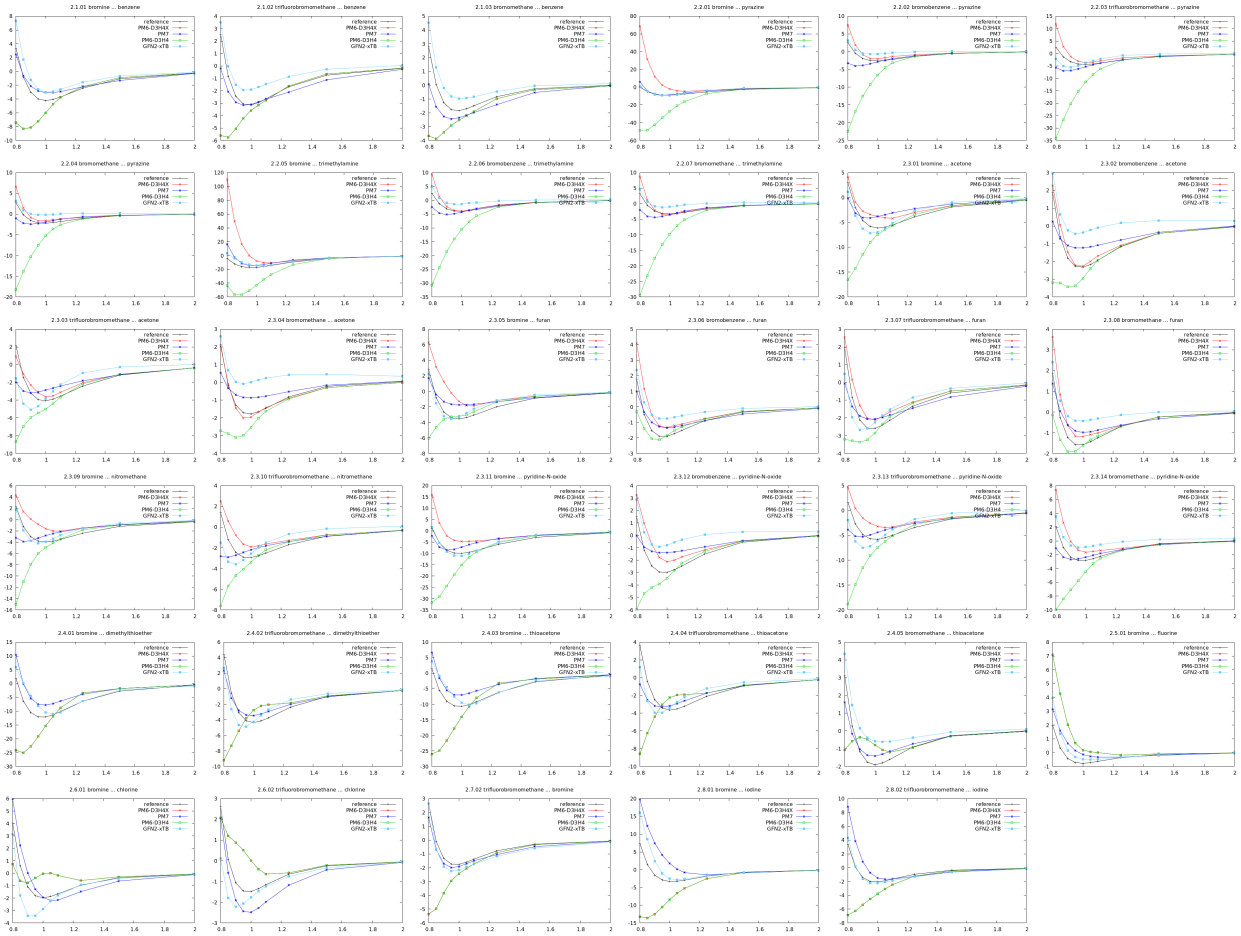

Figure S6: SH250×10 dissociation curves computed with semiempirical QM methods - iodine

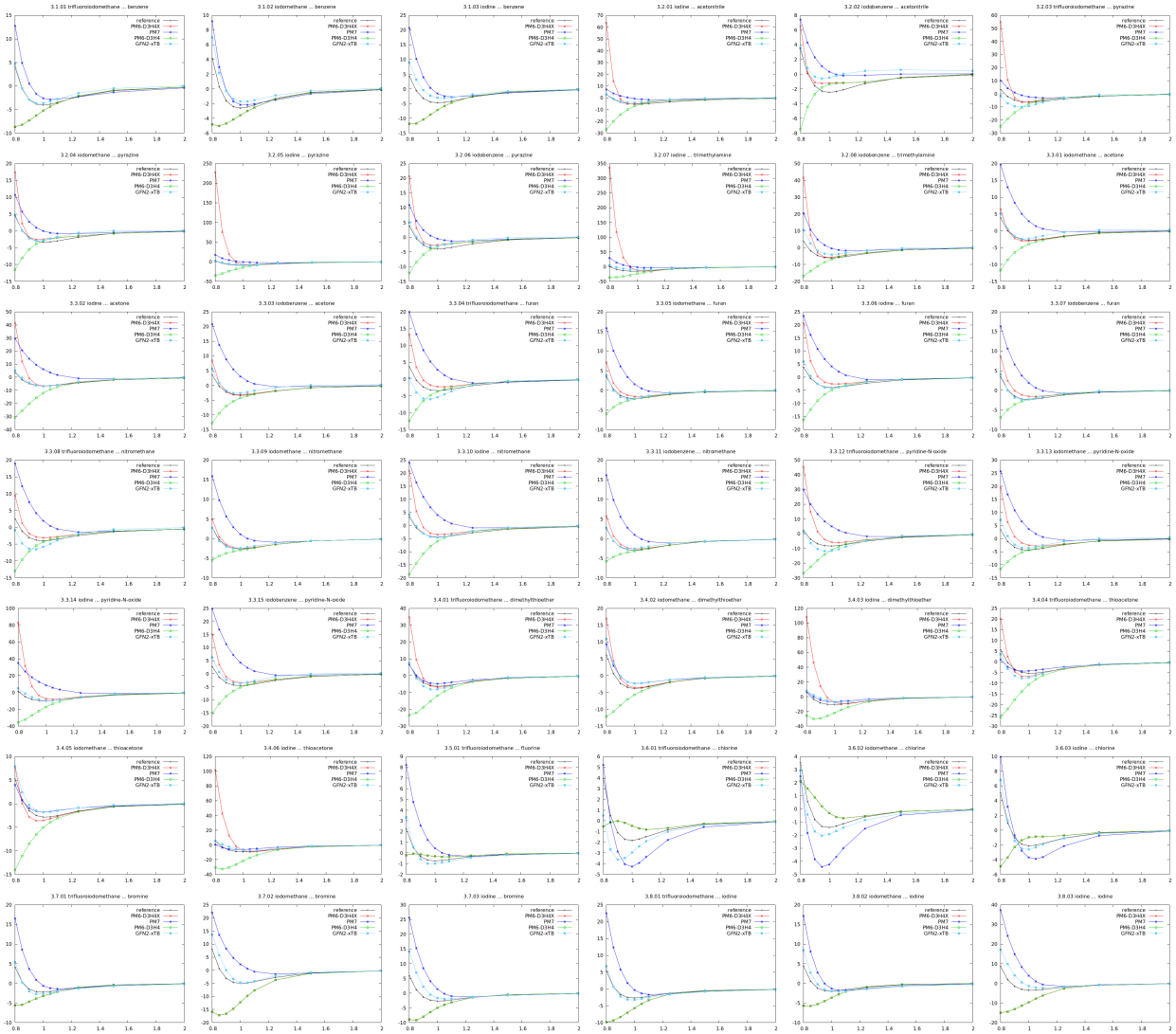

Figure S7: SH250×10 dissociation curves computed with semiempirical QM methods - sulfur

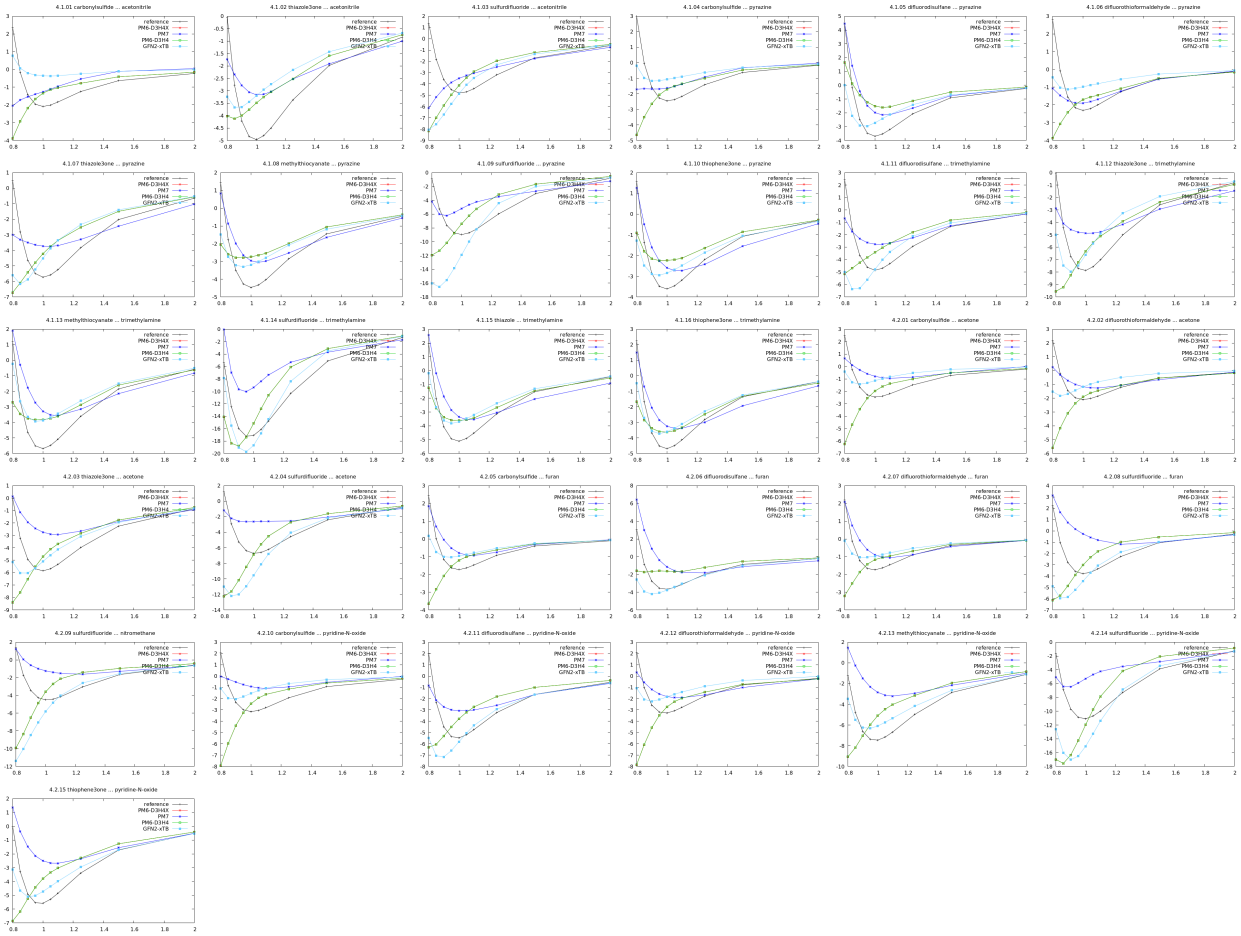

Figure S8: SH250×10 dissociation curves computed with semiempirical QM methods - selenium

Figure S9: SH250×10 dissociation curves computed with semiempirical QM methods - phosphorus

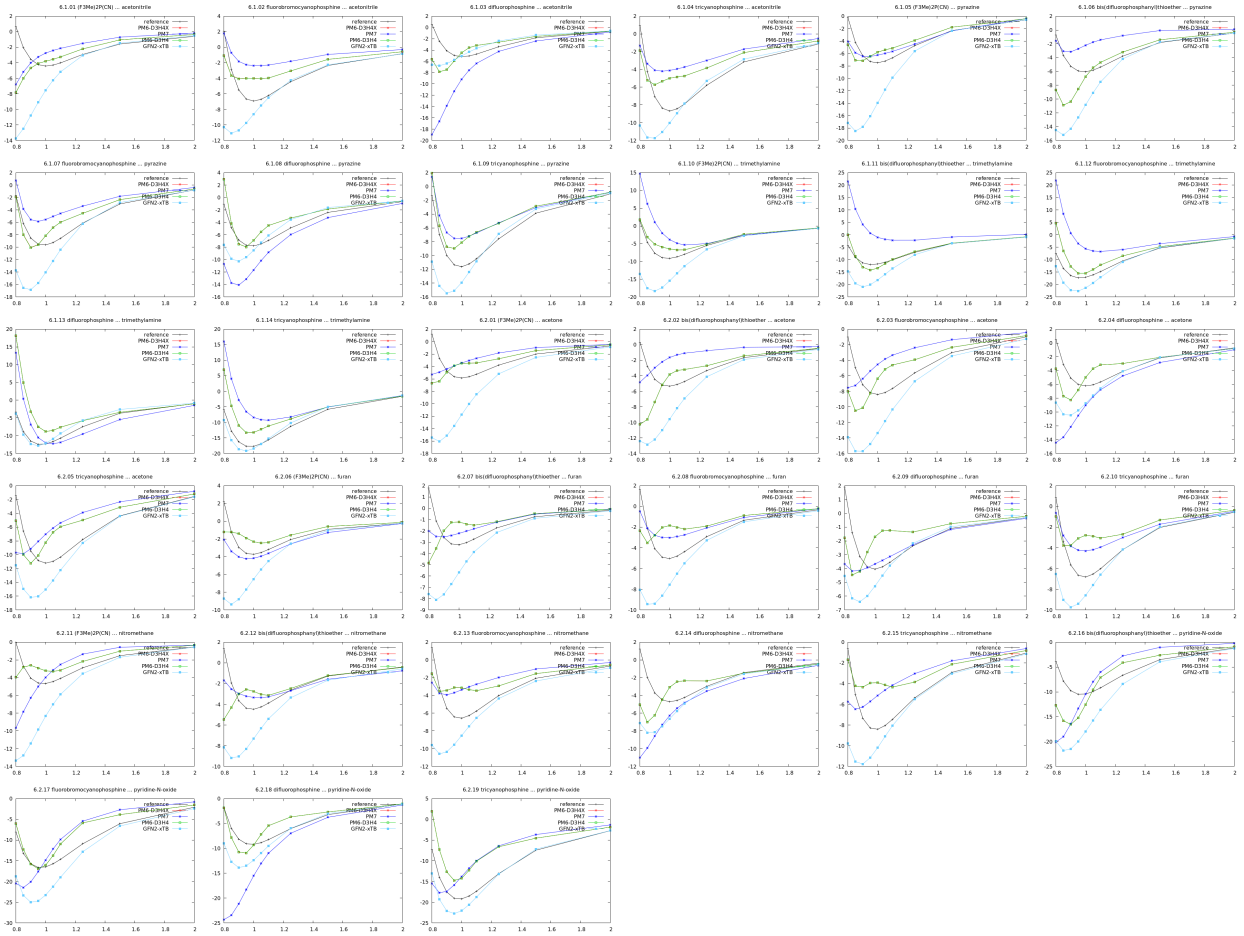

Figure S10: SH250×10 dissociation curves computed with semiempirical QM methods - arsenic